Introduction to R - Tasks

Green-Lab

2023-09-29

Task 1: Create your R intro folder where you have to create an R-Markdown file through your R-Studio IDE and make a nested folder named "data" and paste your dataset "aggregated_results.csv".

Task 2: Create an R-script or R-Markdown file.

Task 3: Install and import "tidyverse" library.

Task 4: Search for "tidyverse" and "dplyr" cheat sheet.

Task 5: Import the csv file named "aggregated_results.csv" on a dataframe named "df".

Task 6: Inspect the first 10 rows (head) and last 10 rows (tail) of the data.

Task 7: Inspect the first/last 10 rows of 2 specific columns of your interest.

Task 8: Rename "Loading_time", "Memory.Usage..KB.", "GPU.Load..." to "loading", "mem" and "gpu" respectively.

Task 9: Make a new dataframe and pass from the original dataframe (df), the columns: "Energy_J", "n_bytes", "loading", "browser", "gpu", "mem" filtering out loading values that exceed 8000. In this new dataframe, you will proceed to the next tasks.

Task 10: Use the Sample function for a new column named "device" to the new dataframe with a set of 3 devices of your choice.

Task 11: Make "browser" and the "device" a factor using mutate.

Task 12: Normalize the "Energy_J" column values (log transform) pass them in a new column named Energy_J_log and do the same for its square root.

Task 13: Group per device you sampled and make a summary of the mean, median, and standard deviation.

Task 14: Inspect which of the data frames columns are numeric, pass their names in a char vector and make a histogram for each one (tip: you will need lapply, unlist, and is.numeric).